

## **SUPPORTING STATEMENT PART B: COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS**

### **B.1. RESPONDENT UNIVERSE AND SELECTION METHODS**

Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

#### Overview:

On behalf of the Directorate of Computer and Information Science and Engineering (CISE) of the National Science Foundation (NSF), the Computing Research Association's (CRA) Center for Evaluating the Research Pipeline (CERP) will be conducting the **CISE REU Past Participant Survey - 2021 Impact of REU Participation on Career Pathways** project. The project, which is designed to evaluate the effectiveness of research experiences for undergraduates (REUs) for promoting computing research careers, involves the collection of data from three groups of survey respondents: (1) NSF CISE REU past participants, (2) NSF REU mentors, (3) participants in other undergraduate research experiences).

#### Description and Numerical Estimates:

- (1) NSF REU past participants: This group will be comprised of undergraduate students in the NSF CISE REU Sites or REU Supplements program who participated between 2013 and the current time. NSF records indicate that there are approximately 13,700 unique individuals in this group who will be contacted for the survey. Samples will not be used; all participants will be invited to participate.

(2) NSF REU mentors: This group will be comprised of past or current CISE REU mentors (including PIs, co-PIs, or others in a mentoring role) who have received an NSF CISE REU Site or Supplement project award between 2013 and the current time. NSF records indicate that there are approximately 6,600 unique individuals in this group who will be contacted for the survey. Samples will not be used; all mentors will be invited to participate.

(3) Comparison participants who have had other undergraduate research experiences: CERP will generate a list of potential comparison respondents from other surveys that CERP has administered over the past several years. This will include the following sources.

- a. CERP has collected yearly data since 2013 from undergraduate students in computing via the **Data Buddies Survey (DBS)**. As part of their DBS participation, respondents are asked about the types of research experiences they have had during their undergraduate years. DBS respondents who have reported engaging in an undergraduate REU would be eligible to take part in the current study as a comparison group participant.
- b. CERP has also evaluated several **non-NSF CISE REU programs**, with data collected as far back as 1994. Program participants from 2013 or later would be eligible to take part in the current study as a comparison group participant.

For both types of comparison group participants, only those who have given CERP permission to contact them for future surveys and have provided a contact email address would be recruited to participate in the current study. CERP estimates this

number to be approximately 7,000 individuals. Notably, because the subset of DBS comparison participants to be pulled do not report the formal name of their REU, CERP is unable to determine the full population size for this group. However, CERP will invite all 7,000 individuals with a non-NSF REU from these sources to participate in the current study.

### Expected Response Rates

Expected response rates for the current study have been estimated based on response rates that CERP has achieved in previous survey research projects that, although not connected to the current study, have used similar populations and methods. For example, CERP's DBS data collection with respondents who have agreed to participate in follow-up surveys has had a response rate of approximately 30%. The current study's potential participants are similar in background to the DBS participants; thus, CERP expects a response rate that is at least as high as that. Because the potential respondents in the current study have by definition had a higher-than-average level of engagement in computing research and invested significant time and effort to their research experiences, CERP believes that these respondents' motivation to participate will, on average, be higher than that of respondents in other CERP surveys. Additionally, CERP's recruitment process will include personalized appeals to participate, an emphasis on the respondents' ability to contribute to understanding and improving undergraduate research experiences, and multiple attempts to elicit participation. Efforts to maximize participation rates are further described in Section B.3.

## **B.2. PROCEDURES FOR THE COLLECTION OF INFORMATION**

**Describe the procedures for the collection of information including:**

- **Statistical methodology for stratification and sample selection**
- **Estimation procedure**
- **Degree of accuracy needed for the purpose described in the justification**
- **Unusual problems requiring specialized sampling procedures**
- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden**

### **Overview**

Data collection for this project intends to assess the impact of REU participation on career pathways. The specific research and evaluation objectives are to:

1. Identify the career trajectory of the REU participants since their participation in the REU program including degrees they received, institutions they attended, and their current status (e.g., employed, graduate students).
2. Document the structure of the REU experience that the respondents participated in. These may include the type of REU (e.g., Site, Supplement), location of REU, and timing of REU.
3. Describe the REU mentors' perceptions of the REU program on the student participants and the mentors' career development.
4. Examine the skills the participants gained and experiences they had during their REU participation. These may include technical skills, information on graduate school application process, and research training.
5. Analyze the predictors of computing career choices, specifically focusing on the experiences and participant characteristics that are associated with the participants' interest in and ultimate selection of research careers in computing.

6. Compare the backgrounds, skills, education, and career pathways of NSF REU participants to a comparison group of participants who have engaged in other, non-NSF REU research activities and experiences.

### **Procedure**

Three types of respondents will be recruited for the survey: (1) former NSF CISE REU participants; (2) former (and/or current) NSF CISE REU mentors; (3) former participants in non-NSF REU experiences. The former NSF CISE REU participants and the former participants in non-NSF REU experiences will complete the REU Past Participants Survey, and the NSF REU mentors will complete the REU Mentor Survey. Procedures for recruiting each group are described below.

Former NSF CISE REU participants and mentors: NSF has provided CERP with two separate lists. One list includes the following information for former undergraduate participants in the NSF CISE REU Sites or Supplement programs:

- Participant name
- Role
- Email Address
- NSF Directorate/Division
- Award information (e.g., ID number, title, type, start date, co-PI list, other project dates)
- Institution name (REU and home)
- Participant demographics (gender, ethnicity and race information, “handicapped description”)
- Participation measures (FTE, Person Months Worked)

The second list includes the following information for NSF CISE REU Sites or Supplements mentors:

- Mentor name
- Email address
- NSF Directorate/Division
- Award information (e.g., ID number, title, type, start date, co-PI list, other project dates)
- Institution name
- Mentor demographics (gender, ethnicity and race information, “handicapped description”)

CERP will send a recruitment email (one version for participants, another version for mentors) to all individuals in these lists. Samples will not be drawn.

Former participants in non-NSF REUs: As described previously, CERP will identify respondents from previous research and evaluation projects who: (1) gave permission to be contacted for follow-up; and (2) participated in an undergraduate research experience that was comparable to an NSF CISE REU. CERP will send a recruitment email to all individuals meeting these criteria. Samples will not be drawn.

For all potential respondents, the **initial recruitment email** will introduce to CERP as the research and evaluation arm of the Computing Research Association (CRA), which has a strong reputation within the computing community. The email will include a description of CERP as the organization conducting the survey on behalf of NSF, a description of the survey project and its purpose, an emphasis on respondents’ ability to contribute to future REU improvements through their participation, and a request to complete the one-time online survey. The email will contain a link to the appropriate participant or mentor survey in Qualtrics as well as contact information for CERP that respondents can use to ask additional questions about the survey or their participation in it.

After two weeks have passed, all respondents who have not yet completed a survey will be sent a **reminder email** that once again describes the project and its purpose and invites them to participate by completing the online Qualtrics survey. The reminder will again include contact information for CERP for any questions about the survey.

After another two weeks have passed, respondents who still have not completed a survey will be sent a **final email** with a reminder about the project and a request, with a link, to complete the online Qualtrics survey. CERP contact information will again be included.

Survey invitations will be sent to the full set of potential participants; no samples will be selected. For two of the groups – the past NSF CISE REU participants and mentors – this comprises the full population. For the third group, the full universe of participants is unknown, because a large portion of those recruited will not have reported the name of the REU that they took part in. CERP will attempt to recruit all participants within this group.

- **Statistical methodology for stratification and sample selection**

Not applicable. CERP is not selecting a sample to survey.

- **Estimation procedure**

#### **Respondent-level non-response analyses**

Based on data collection from similar populations, CERP expects the overall response rate for the survey to be above 30%. To understand whether there are systematic differences in who responds to the survey, CERP will conduct and report on respondent-level non-response bias analysis using the data that are available for the full potential respondent population. This analysis will vary depending on the type of respondent.

NSF CISE REU participants and mentors: For both the NSF CISE REU participants and mentor surveys, CERP will examine whether there are different rates of survey participation as a function of CISE division, participation year, REU type, gender, or race/ethnicity. Non-response patterns will be investigated separately for participants and mentors.

Non-NSF REU participants: For the non-NSF REU participants in the comparison group, the available data for conducting an analysis of non-response bias will vary depending on which prior CERP survey(s) the respondent has completed. Basic respondent non-response analyses should be able to be completed using gender, race/ethnicity, and year of REU participation.

### **Item-level non-response analyses**

In addition to respondent-level non-response analyses, concerns have also been raised about non-response bias at the levels of individual survey items.<sup>1</sup> Although non-response bias analyses of all survey items are not feasible, CERP will monitor and report on survey items that are unanswered by a large percentage of respondents overall and within each of the three respondent groups.

- **Degree of accuracy needed for the purpose described in the justification**

Assuming a 30% response rate, CERP expects to receive 4,110 surveys from past NSF REU participants, 1,980 surveys from NSF REU mentors, and 2,100 surveys from non-NSF REU participants. The overall N of 6,210 NSF and non-NSF REU participants is adequate for the reporting of the descriptive information that is detailed in [Objectives 1, 2, and 4](#). The expected N of 1,980 in the mentor survey is adequate for achieving [Objective 3](#), which is a descriptive summary of mentors' program perceptions. The separate N's of 4,110 and 2,100, for NSF and

---

<sup>1</sup> Baker, R., Brick, J. M., Keeter, S., Biemer, P., Kennedy, C., Kreuter, F., ... & Terhanian, G. (2016). Evaluating survey quality in today's complex environment. *American Association for Public Opinion Research (AAPOR)*.



non-NSF REU participants, respectively, is adequate for the comparative analyses described in [Objective 6](#). The overall participant N of 6,210 should be adequate for analyses examining most of the potential predictors of choosing computing careers ([Objective 5](#)); however, even with the large overall N, there may be small numbers of respondents for some subgroups of interest (such as certain racial/ethnic groups) that may limit CERP's ability to include those variables in any predictive analyses. The research team will conduct power analyses during the reporting stage and provide interpretation of the findings in that context.

- **Unusual problems requiring specialized sampling procedures**

Not applicable. There are no specialized sampling procedures being used.

- **Any use of periodic (less frequent than annual) data collection cycles to reduce burden**

Not applicable. There is only one planned survey administration for this project.

### **B.3. METHODS TO MAXIMIZE RESPONSE RATES AND THE ISSUE OF NONRESPONSE**

**Describe methods to maximize response rates and to deal with issues of nonresponse. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.**

CERP will take several measures to maximize response rates for the participant and mentor surveys.

#### **Quality of contact information:**

CERP will have access to contact information provided by NSF for two of the three respondent types (former NSF CISE REU participants and mentors). For the third respondent type (former non-NSF CISE REU participants who have given permission to be contacted), respondent contact information will come from CERP's own records and will have been provided by the respondents themselves. Thus, we expect the quality of the contact

information to be better than average. Additionally, for emails that bounce back after the first recruitment contact, CERP will send survey invitations to any updated email addresses that are included in email bouncebacks, as available. CERP will attempt to find updated email contact information, as possible, using searches of publicly available data (such as university websites and internet search engines).

**Multiple recruitment prompts:**

As described in Section B.2, there will be multiple recruitment invitations sent to respondents to encourage high participation rates. All respondents will receive an initial individualized recruitment email with study information and a link to complete the online survey. After approximately two weeks, a reminder email will be sent to those who have not yet completed the online survey. After another two weeks has passed, a final reminder will be sent to those who still have not completed the online survey. CERP will evaluate the response rates two weeks after the final reminder to determine whether to extend the data collection window further; although surveys that are in the field longer have higher response rates, this must be balanced against the need for timely completion of data collection.<sup>2</sup>

**Survey Convenience:**

Respondents will be able to easily link to the online Qualtrics survey – via computer, tablet, or mobile device – through a personalized survey link provided in the recruitment email sent by CERP. Both the participant and mentor surveys have been designed to minimize respondent burden wherever possible. For example, the overall survey completion time should be 20 minutes at most and may take less time for many respondents. Skip patterns and survey

---

<sup>2</sup> Baker, R., Brick, J. M., Keeter, S., Biemer, P., Kennedy, C., Kreuter, F., ... & Terhanian, G. (2016). Evaluating survey quality in today's complex environment. *American Association for Public Opinion Research (AAPOR)*.

flow have been carefully considered in the survey design to ensure that respondents are presented with only the items that are relevant to them. Respondents will be able to start and pause the survey so that they can complete it at their convenience.

**Information, motivation, and resources for respondents:**

Communications with potential respondents will promote participation in several ways. Emails sent to potential respondents will include information about the study and its purposes, thereby ensuring that they are fully informed about the project. The emails will also underscore the importance of their role in contributing to improving the knowledge base about undergraduate research experiences and their impacts. Furthermore, as part of the Computing Research Association (CRA), CERP will be able to leverage CRA's strong reputation within the computing community to promote high levels of survey participation. All emails will include an introduction to CERP as the research and evaluation arm of the CRA. The emails will explain CRA's role in administering the survey on behalf of NSF. The emails will include a link to more information about CRA and CERP, as well as contact information for CERP staff for those seeking more information about any aspects of the study.

**Non-response bias analysis:**

CERP will use available data from the universe of potential participants to determine whether subgroups of respondents are participating in the survey at different rates. CERP will also examine whether individual survey items have high rates of missing responses. Both of these non-response analyses are described further in Section B.2, and both will be included in the reporting of the study results.

#### **B.4. TESTS OF PROCEDURES**

**Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.**

##### **Background:**

Procedures: CERP has used this study's planned recruitment procedures in dozens of previously survey research projects. These procedures have been tested and refined repeatedly to enhance efficiency and improve response rates.

Instrument development: Many of the items in the Past REU Participants Survey and the REU Mentor Survey were initially developed and implemented in CERP's annual Data Buddies Survey or in previous evaluations of undergraduate research programs conducted by CERP. Consequently, many items in the current instruments have been tested and refined over a period of many years. Of course, new items have also been added to these surveys, and some previously developed items have been updated to enhance their applicability to the current study purposes. For both the previously-developed and new survey items, community stakeholders and subject matter experts in REUs have been recruited to contribute to item development and refinement. CERP will also engage in significant instrument testing prior to commencing data collection, as described below.

##### **Testing plans:**

Both the Past REU Participants Survey and the REU Mentor Survey will be programmed in Qualtrics. CERP will conduct comprehensive testing of the survey, including the following:

- o Proofing survey question text

- o Confirming appropriate question types have been programmed (e.g., multiple versus single response options)
- o Ensuring that the survey skip logic has been appropriately programmed
- o Checking that the survey functions adequately in multiple browsers
- o Checking that the order of survey questions facilitates easy comprehension and “flow” for respondents; i.e., that the order and placement of questions feels natural and flows well
- o Pilot testing the survey will be conducted in two ways. First, internal CRA staff who are not familiar with the study will be asked to complete the survey and provide feedback on the clarity of items and the overall survey administration experience, including the amount of time it takes them to complete it. After that process is complete, if additional testing is warranted, pilot testing will be conducted with fewer than 10 individuals outside of the CRA organization.

#### **B.5. CONSULTANTS**

**Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

NSF has contracted with the Computing Research Association’s (CRA) Center for Evaluating the Research Pipeline (CERP) to design and conduct this survey project. Names, titles, contact information, and roles of the CERP staff who will be involved in the study are described below.

- o **Dr. Burcin Tamer, Director of CERP (202-266-2935)**

Tamer will have general oversight over the project and manage the contract overall. She has led the study design (including statistical aspects of the design) and instrument development.

o **Heather Wright, Associate Director of CERP (202-266-2945)**

Wright will be responsible for project management and staff oversight. She has contributed to the study design (including statistical aspects of the design) and led instrument development.

o **Kristi Kelly, Research Associate (202-266-2935)**

Kelly will be actively involved in data management, analysis, and reporting activities. She has contributed to the study design (including statistical aspects of the design) and instrument development.

o **Taniya Ross-Dunmore, Research Assistant (202-266-2935)**

Ross-Dunmore will be engaged in administrative aspects of the project, including programming and testing the online surveys.